

## Комплексный алгоритм вычисления оптического потока с использованием взвешенной меры подобия

С. В. ПАНИН<sup>1,2,\*</sup>, В. О. ЧЕМЕЗОВ<sup>1</sup>, П. С. ЛЮБУТИН<sup>1</sup>

<sup>1</sup>Институт физики прочности и материаловедения СО РАН, Томск, Россия

<sup>2</sup>Национальный исследовательский Томский политехнический университет, Россия

\*Контактный e-mail: svp@ispms.tsc.ru

Предложен и протестирован алгоритм вычисления оптического потока для пар изображений, ориентированный на повышение точности определения перемещений на границах объектов. В его основе лежит комбинация таких методов, как: а) блочное сопоставление; б) расчет меры подобия блоков; в) анализ нескольких каналов информации; г) иерархический подход (пирамиды изображений); д) корректировка распределения функции подобия блоков изображения на основе весов билатерального фильтра; е) субпиксельное уточнение смещений путем аппроксимации функции подобия блоков изображения вблизи максимума соответствующего распределения; ж) сглаживание вычисленного поля оптического потока.

Результаты экспериментального исследования показали, что применение предложенного алгоритма является перспективным для задач стереозрения. Последовательное изменение параметров предложенного алгоритма позволило определить их рациональные значения, обеспечивающие снижение средней абсолютной ошибки определения конечной точки.

*Ключевые слова:* обработка изображений, оптический поток, оценка движения, сопоставление областей, блочное сопоставление, границы объектов, ошибка определения конечной точки, взвешенная мера подобия, комплексный алгоритм, субпиксельная точность.

### Введение

Компьютерное зрение получает все большее распространение в различных областях исследований и промышленности. Одним из актуальных направлений применения машинного зрения является решение задач экспериментальной механики, в частности оценка деформации по анализу последовательностей изображений или видеопотока. В рамках такого подхода используют анализ как двумерной информации и соответственно оценивают распределение деформации в плоскости наблюдения, так и трехмерных сцен на основе карт глубин с последующим определением перемещений по трем координатам. На рынке аппаратно-программных средств, принцип действия которых основан на методе корреляции цифровых изображений (Digital Image Correlation — DIC), получили широкое распространение системы VIC 3D [1], Strain Master [2] и др. Однако компьютерные коды в рамках их программного обеспечения являются закрытыми, что

не позволяет гибко подходить к решению ряда прикладных задач. Таким образом, разработка методов оценки перемещений в системах трехмерного зрения — актуальная научно-техническая задача.

Традиционным подходом к оценке перемещений, в том числе в системах стереозрения, является определение оптического потока. Расчет оптического потока — одна из наиболее актуальных проблем в области компьютерного зрения. Обзор соответствующих методов для решения данной задачи приведен в [3]. Показано, что большая часть алгоритмов основана на использовании оригинального метода, предложенного в работе Хорна и Шанка [4], в которой введено понятие регуляризации — параметра, определяющего (регулирующего) выбор каждого вектора оптического потока в зависимости от степени его подобия соседним векторам. Данный параметр может быть представлен как

$$E = E_{data}(\bar{u}) + \lambda E_{reg}(\bar{u}), \quad (1)$$

где  $\bar{u} = (u, v)$  — поле оптического потока;  $E_{data}$  — мера подобия блоков изображений (data cost);  $E_{reg}$  — мера подобия векторов в оптическом потоке;  $\lambda$  — регуляризационный коэффициент.

Ключевым критерием, характеризующим эффективность работы трехмерной системы оценки деформации методом корреляции цифровых изображений, является точность (включая помехоустойчивость) определения оптического потока. Однако такие системы, как правило, должны работать в режиме реального времени [1, 2], поэтому требуется снижение вычислительных затрат, что может быть достигнуто за счет применения более простых алгоритмов расчета перемещений либо использования аппарата параллельных вычислений [5].

Анализ данных тестирования различных алгоритмов определения оптического потока, представленных в публичной базе MiddleBury benchmark [6, 7], показал, что большая их часть основана на блочном сопоставлении [8], так как этот подход обеспечивает возможность обрабатывать разные фрагменты изображения независимо друг от друга, что позволяет уменьшать вероятность ошибочного определения смещений [3]. Проведен детальный анализ нескольких алгоритмов, характеризующихся наименьшей среднеквадратичной ошибкой определения смещений, а также обладающих наименьшими временными затратами. Показано, что наилучшей точностью оценки перемещений обладает алгоритм NNF (Nearest Neighbor Field), основанный на определении поля ближайших соседей [9], а также алгоритм OFLAF (Optical Flow via Locally Adaptive Fusion), использующий в качестве критерия определения перемещения минимум объединенных локальных мер подобия, рассчитанных по нескольким каналам данных (трем цветовым каналам, полутоновому каналу яркости, каналам градиентов и т. д.) [10, 11]. С точки зрения обеспечения максимального быстродействия следует выделить алгоритм SF (Simple Flow) [12], основанный на поиске минимума меры подобия с применением корректировочных весовых коэффициентов. Последний доступен в виде открытого кода в публичной библиотеке OpenCV [13].

Поскольку в работах [9–12] не приведено описание значений параметров и коэффициентов, используемых, например, для выполнения фильтрации в вышеперечисленных алгоритмах, а также с целью достижения максимальной точности определения оптического потока при приемлемых вычислительных затратах, в настоящей работе поставлена задача разработки гибридного алгоритма, использующего преимущества подходов, примененных в каждом из них, а также подбора параметров их эффективного функционирования.

## 1. Описание алгоритма

Для вычисления оптического потока могут быть использованы следующие подходы:

- поиск смещений методом сопоставления блоков (PatchMatch);
- анализ нескольких каналов информации;
- иерархический подход (пирамиды изображений), использующий переход от грубого к мелкому масштабу рассмотрения изображений (coarse-to-fine);
- корректировка функции подобия блоков изображения на основе весов билатерального фильтра;
- субпиксельное уточнение смещений путем аппроксимации функции подобия блоков изображения вблизи максимума;
- сглаживание вычисленного поля оптического потока.

В работе использованы обозначения:  $I_t$  — изображение в начальный момент времени,  $I_{t+1}$  — изображение в текущий (следующий) момент времени,  $I(\bar{i})$  — яркость пикселя в точке  $\bar{i} = (i, j)$ ,  $\bar{m} = (m, n)$  — радиус окрестности пикселя (рис. 1).

**Блочное представление изображения.** Невозможность точно определить смещение одного пикселя связана с недостатком локальной информации. Предполагаем, что на соседних кадрах в некоторой окрестности пикселя яркость блока изображения должна оставаться постоянной. В литературе это получило наименование *ограничение постоянства яркости* (brightness constancy). В процессе построения вектора оптического потока для пикселя  $I(\bar{i})$  производится сопоставление блоков (участков изображения)  $\zeta(I(\bar{i}), \bar{m})$  малой окрестности этого пикселя  $I(\bar{i})$  радиуса  $m$ , для которой данный элемент изображения  $I(\bar{i})$  является центральным. Пример такой области показан на рис. 1. В работе радиус блока выбран равным  $\bar{m} = (5, 5)$  [14] как наиболее эффективный, что будет доказано экспериментально (см. разд. 3).

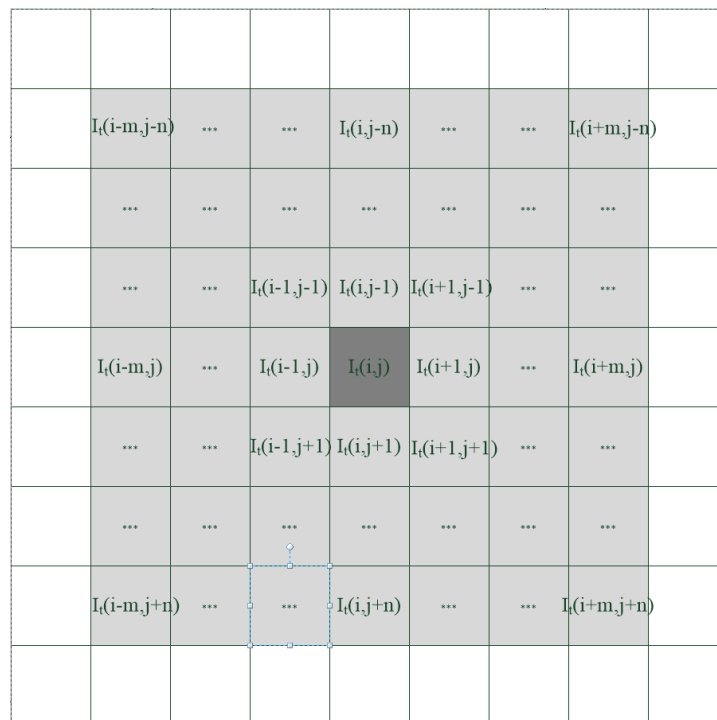


Рис. 1. Представление окрестности пикселя в виде блока изображения

**Оценка соответствия блоков изображения.** Для получения оценки, характеризующей соответствие блоков сравниваемых изображений, проводится расчет функции вида

$$E(\bar{u}) = \sum_{i,j} [\rho(I_t(i, j) - I_{t+1}(i + u, j + v))], \quad (2)$$

где  $\rho(x) = \exp(-|x|)$ . Использование такой функции позволяет перейти от классического подхода, основанного на поиске максимума меры подобия в пределах области поиска (например, максимума функции корреляции), к поиску максимальной вероятности смещения [15]. При этом в рамках расчета функции меры подобия удастся увеличить вклад от наиболее похожих элементов в области поиска и уменьшить вклад от наименее подобных. При данном подходе функция, обеспечивающая выбор наилучшего соответствия (2), имеет вид

$$\bar{u} = \arg \max_{\bar{u}} E. \quad (3)$$

**Анализ нескольких каналов информации.** Традиционно при сопоставлении кадров используется от одного (для черно-белых) до трех каналов информации (RGB — для цветных изображений). В [8] показано, что можно использовать одновременно до 31 канала данных, количество каналов, вообще говоря, может быть и больше. В этой же работе представлен график зависимости точности определения оптического потока от количества каналов анализируемых данных. Показано, что эта зависимость выходит на насыщение уже при числе каналов выше восьми. В настоящей работе для вычисления оптического потока при обработке полутоновых изображений предложено использовать три канала данных: один полутоновой канал самого изображения и два канала, характеризующих градиент яркости, которые были получены с помощью оператора Собеля [16] в вертикальном и горизонтальном направлении.

**Иерархический подход.** Метод вычисления оптического потока “coarse-to-fine” (от грубого к точному), предложенный в [17], также широко применяется при обработке изображений и видеопотоков, поскольку позволяет за меньшее количество итераций добиться приемлемого уровня точности определения смещений. Метод обеспечивает два конкурентных преимущества: возможность определения больших смещений и увеличение скорости построения оптического потока. При этом на каждой итерации последовательно анализируются слои (уровни) пирамиды изображений, начиная от самого низкого (изображения минимального размера).

**Расчет взвешенной функции подобия блоков.** Данный подход позволяет при построении оптического потока корректно (с наилучшей устойчивостью) определить перемещения на границах объектов на изображении. Эта проблема особо актуальна для систем стереозрения, поскольку при съемке рабочей сцены двумя камерами, разнесенными в пространстве, часто возникают ошибки на границах объектов, что вызвано их взаимным перекрытием. Для решения этой задачи используется кроссбилатеральная фильтрация функции подобия [18], коэффициенты которой рассчитаны по входному (исходному) изображению. Функция подобия блоков, учитывающая весовые коэффициенты, имеет следующий вид:

$$E(\bar{u}) = \sum_{i,j} w_c w_d [\rho(I_t(i, j) - I_{t+1}(i + u, j + v))],$$

где  $w_c = \exp\left(-\frac{(I_t(\bar{i}) - I_t(\bar{i} + \Delta\bar{i}))^2}{2\sigma_c^2}\right)$  — весовая функция для данных о цвете;

$w_d = \exp\left(-\frac{(\Delta \bar{i})^2}{2\sigma_d^2}\right)$  — весовая функция для данных о расстоянии. Обе весовые функции вычисляются внутри блока изображения. На рис. 2, б приведен пример распределения весовых коэффициентов, рассчитанных для центрального пикселя блока изображения, представленного на рис. 2, а. Видно, что для этого пикселя наибольший вес (рис. 2, б) имеют пиксели с примерно тем же уровнем яркости, а также расположенные ближе к нему. Шкала значений весовых коэффициентов приведена на рис. 2, в. В работе экспериментально подобраны значения коэффициентов весовых функций:  $\sigma_c = 90$  и  $\sigma_d = 16$ .

**Субпиксельное уточнение.** Повышение точности определения перемещений в оптическом потоке может быть обеспечено за счет субпиксельного разрешения. Простым и быстрым способом достижения субпиксельной точности определения перемещений является сглаживание смещений ближайших соседей с помощью полиномиальной аппроксимации с последующим поиском ее локального максимума [19]. Предположим, что мера подобия для блока изображения в оптическом потоке по каждой из осей имеет параболическое распределение. Для ее аппроксимации используются величины перемещений, рассчитанные для каждого пикселя на основании значений его четырехсвязной окрестности. Таким образом, смещение может быть уточнено с субпиксельной точностью относительно определенного ранее целочисленного значения:

$$\Delta i = \frac{E(u-1, v) - E(u+1, v)}{2E(u-1, v) - 4E(u, v) + 2E(u+1, v)},$$

$$\Delta j = \frac{E(u, v-1) - E(u, v+1)}{2E(u, v-1) - 4E(u, v) + 2E(u, v+1)}.$$

**Сглаживание оптического потока.** При использовании иерархического подхода возникает множество ошибок определения перемещений, так как при уменьшении изображения теряются его детали, а сама оценка меры соответствия “загрубляется”. Чтобы ошибки при определении перемещений, возникающие на “нижних” слоях пирамиды (при малом разрешении), не “переносились” далее на анализ “верхних” слоев, необходимо проводить фильтрацию оптического потока. В [20] показано, что для решения этой задачи эффективно применение медианного фильтра с апертурой  $5 \times 5$  пикселей. Помимо решения задачи фильтрации явных ошибок, сглаживание исходного изображения

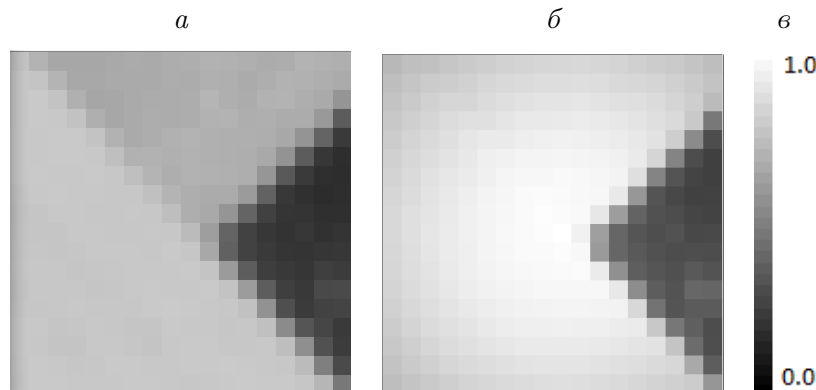


Рис. 2. К пояснению расчета весовых коэффициентов: *a* — исходное (входное) изображение; *b* — карта произведений весовых коэффициентов  $w_c(\bar{i}) \times w_d(\bar{i})$ ; *v* — шкала весовых коэффициентов

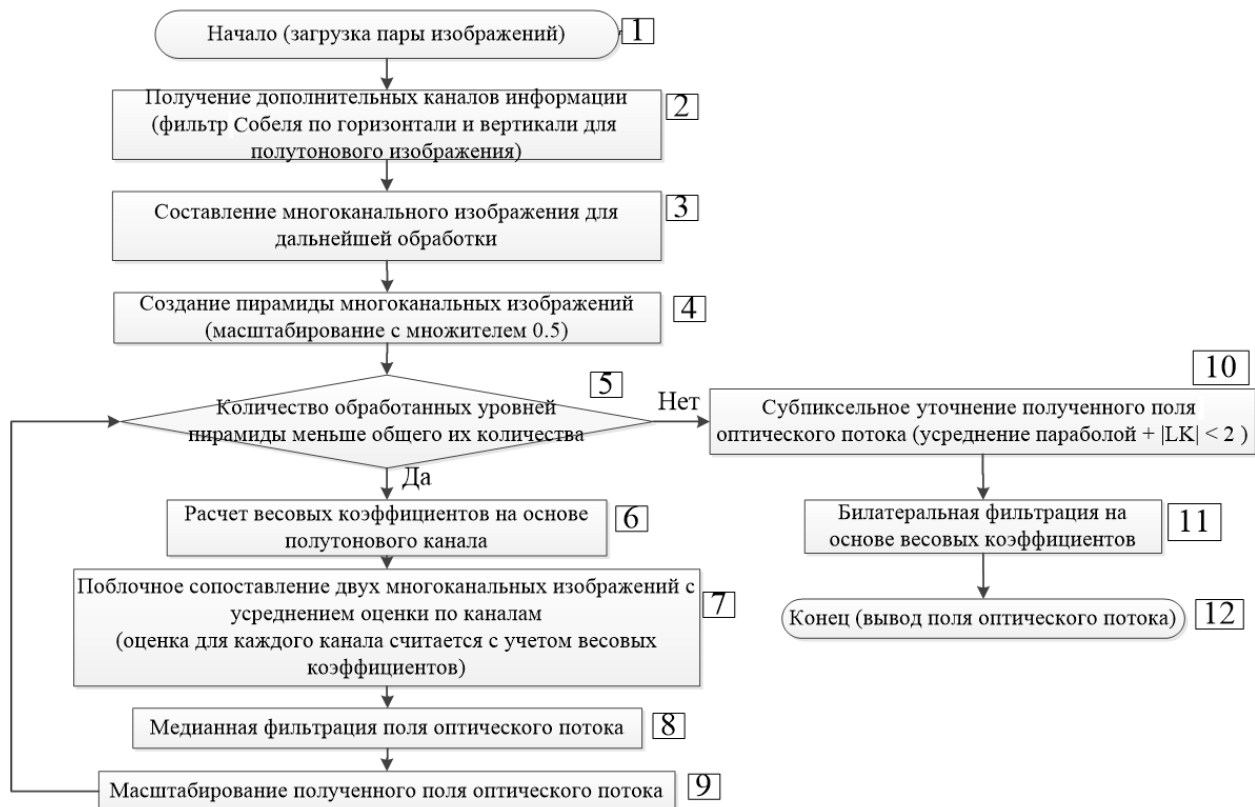


Рис. 3. Блок-схема комплексного алгоритма определения оптического потока на основе взвешенной меры подобия

на основе билатеральных коэффициентов (кроссбилатеральная фильтрация [18]) позволяет более точно определить перемещения на границах объектов в оптическом потоке. Параметры кроссбилатеральной фильтрации рассчитываются по формулам, приведенным выше, отдельно для вертикальной и горизонтальной составляющих оптического потока. Экспериментально были подобраны значения  $\sigma_{fc} = 26$  и  $\sigma_{fd} = 32$  при размере апертуры фильтра  $21 \times 21$ .

Для реализации вышеописанных подходов был предложен единый алгоритм CWFS (Complex Weighted Full Search), включающий использование всех перечисленных методов. Этапы алгоритма (рис. 3):

- формирование входного трехканального изображения (блоки 2 и 3);
- вычисление количества уровней пирамиды в зависимости от входного разрешения и масштабирование входного изображения (блок 4);
- вычисление весовых коэффициентов функции соответствия блоков по входному изображению для каждого уровня пирамиды (блок 6);
- блочное сопоставление, усредненное по каналам для каждого уровня пирамиды (блок 7);
- сглаживание полученного потока для каждого уровня пирамиды (блок 8);
- масштабирование полученного потока и передача его на следующий уровень пирамиды (блок 9);
- субпиксельное уточнение смещений и кроссбилатеральная фильтрация потока на последнем уровне пирамиды (блоки 10 и 11).

## 2. Методика тестирования

Проведено сравнение результатов определения оптического потока, полученных с использованием предложенного в работе алгоритма CWFS (основанного на применении взвешенной меры подобия), с алгоритмами, представленными в публичной тестовой базе MiddleBury benchmark [3, 4]. Следует отметить, что большая часть пар изображений в этой базе характеризуется малыми перемещениями фрагментов на них. Оценку проводили по восьми парам изображений, для которых известны величины модельных смещений. Эффективность алгоритмов вычисления оптического потока с позиции точности и помехоустойчивости оценивали путем расчета средней абсолютной ошибки определения конечной точки (Average Endpoint Error — AEE [4]):

$$AEE = \frac{1}{hw} \sum_{y=1}^h \sum_{x=1}^w \sqrt{(u - u_{GT})^2 + (v - v_{GT})^2}, \quad (4)$$

где  $h, w$  — высота и ширина изображения;  $(u, v)$  — рассчитанный вектор потока;  $(u_{GT}, v_{GT})$  — действительный вектор потока.

В системах стереозрения наибольший интерес представляет повышение точности и устойчивости определения оптического потока на границах объектов [6, 7]. Так, для каждой пары тестовых изображений (рис. 4, *а*) с помощью алгоритма выделения границ Канны (использована его реализация, взятая из публичной библиотеки OpenCV [13] с величинами порогов  $T1 = 100$  и  $T2 = 300$ ) получены изображения с границами объектов (рис. 4, *б*). Для каждой пары изображений получено также поле оптического потока (рис. 4, *в*), традиционно визуализируемое с помощью цветовой шкалы (рис. 4, *г*). Эти изображения далее использованы в качестве маски для расчета параметра AEE (4). Дополнительно по величине данного критерия (с целью повышения наглядности полученных результатов) проведен сравнительный анализ выбранных алгоритмов при обработке полных изображений.

## 3. Результаты тестирования

Результаты тестирования алгоритмов, основанные на расчете AEE, приведены в табл. 1. Первоначально выполнено сравнение предложенного алгоритма CWFS (см. рис. 3) с “реперными”: Classic NL [21] и Simple Flow [12]. Параметры реперных алгоритмов использованы по умолчанию, так как являются наилучшими, по мнению их авторов.

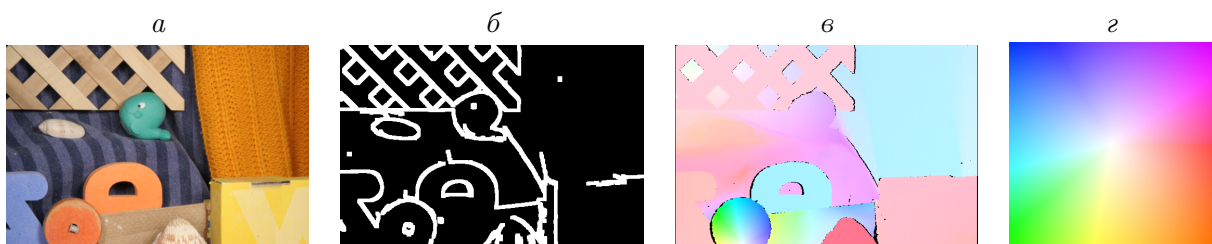


Рис. 4. К иллюстрации способа обработки данных: *а* — исходное изображение; *б* — выделенные границы; *в* — модельный (из базы MiddleBury benchmark) оптический поток; *г* — цветовая кодировка векторов [6, 7]

Т а б л и ц а 1. Средняя ошибка определения конечной точки  $AEE$  при анализе границ объектов / полного изображения

Тип изображения	Алгоритм		
	Classic_elementNL	Simple Flow	CWFS
RubberWhale	0.265 / 0.143	0.431 / 0.291	0.317 / 0.194
Dimetrodon	0.192 / 0.272	0.330 / 0.326	0.197 / 0.247
Venus	0.450 / 0.410	0.385 / 0.418	0.363 / 0.339
Hydrangea	0.410 / 0.207	0.608 / 0.309	0.532 / 0.251
Grove 2	0.308 / 0.270	0.484 / 0.430	0.364 / 0.296
Grove 3	0.982 / 0.808	1.057 / 0.844	0.985 / 0.761
Urban 2	1.583 / 0.863	1.548 / 0.928	2.139 / 0.980
Urban 3	1.191 / 1.033	4.295 / 1.903	3.206 / 1.713
Среднее (по всем типам изображений)	0.673 / 0.501	1.142 / 0.681	1.013 / 0.598

В табл. 1 приведены результаты, полученные при определении оптического потока по всему изображению и только на границах объектов. Картины распределения оптического потока для всех использованных изображений и алгоритмов представлены на рис. 5.

Из сравнения приведенных в табл. 1 данных следует, что предложенный в работе алгоритм CWFS в среднем (см. последнюю строку в каждой графе) обеспечивает лучшее качество построения оптического потока (оцениваемое субъективно при сравнении их картин) и меньшую  $AEE$  по сравнению с Simple Flow. В то же время, судя по величине параметра  $AEE$ , алгоритм CWFS уступает Classic NL как при оценке полного изображения, так и при учете только границ объектов. Несмотря на это, на представленных картинах распределения оптического потока видно, что алгоритм CWFS, равно как и Simple Flow, позволяет получить более “гладкие” перемещения в пределах отдельных объектов и обеспечивает сохранение более четко выраженных границ, в то время как использование алгоритма Classic\_elementNL дает более “зашумленную” картину оптического потока, связанную с локальными вариациями смещений (вследствие ошибок при их определении). Таким образом, предложенный алгоритм CWFS по результатам обработки большинства серий тестовых изображений может быть рекомендован как обеспечивающий высокую точность и гладкость построения оптического потока.

Рассмотрим полученные результаты с учетом специфики обрабатываемых изображений. Величину  $AEE$  анализировали, прежде всего, для случая учета только границ объектов (табл. 1). Однако для сравнения далее в скобках приведено значение данного параметра при рассмотрении всего изображения. Первые три пары изображений получены с помощью видео- и фотокамер и содержат скрытую текстуру (hidden texture). Изображения серии Venus являются стереопарой.

Для серии RubberWhale характерно наличие значительного количества объектов, достаточно равномерно распределенных по изображению, поэтому фон составляет менее 30 % от него. Границы объектов при этом (рис. 5, б) составляют заметную часть изображения. В верхней его части расположена деревянная решетка; алгоритм Classic NL не позволяет точно определить смещения вдоль ее направляющих (рис. 5, д), в то время как для Simple Flow такой недостаток не характерен (рис. 5, е), однако в центре цилиндра (в нижней части изображения) происходит некоторое искажение потока. В целом можно констатировать, что алгоритмы Simple Flow и CWFS дают визуально



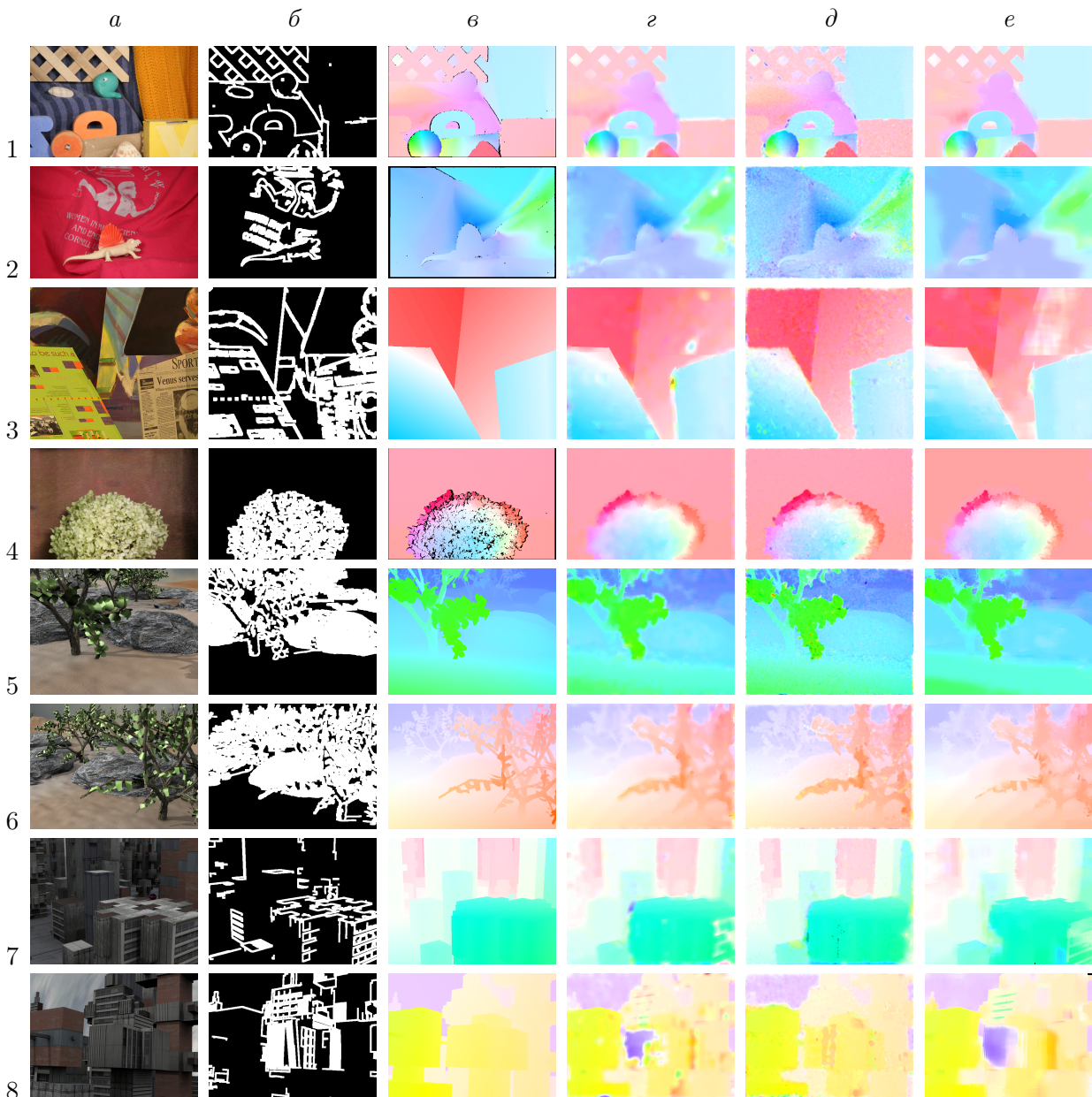


Рис. 5. Тестовые изображения и построенные для них картины оптического потока: *a* — исходное изображение; *b* — выделенные границы; *в* — модельный (истинный) поток; результаты расчета по алгоритмам: CWFS (*г*), Classic NL (*д*), Simple Flow (*е*)

подобные картины распределения оптического потока. Несмотря на это, согласно данным расчета параметра АЕЕ наименьшую ошибку показал Classic NL — 0.265 (0.143), чуть выше она для CWFS — 0.317 (0.194), в то время как максимум соответствует Simple Flow — 0.431 (0.291). В целом ошибка определения оптического потока при анализе только границ для всех исследованных алгоритмов возрастает в 1.48–1.85 раза.

Для пары изображений Dimetradon, наоборот, характерен лишь один объект на довольно однородном фоне, а доля границ объектов (рис. 5, *б, в*) — невелика. Применение алгоритма Simple Flow позволяет достаточно достоверно определять перемещения на границах объектов (рис. 5, *е*), однако это не сопровождается снижением величины АЕЕ,

которая имеет максимум среди трех использованных алгоритмов  $AEE = 0.330$  (0.326). С другой стороны, величина параметра  $AEE$  при построении оптического потока алгоритмом Classic NL имеет минимальное значение —  $0.192$  (0.272), несмотря на то, что поле потока выглядит “зашумленным” (рис. 5, *д*). Для разработанного CWFS величина ошибки лишь чуть выше —  $AEE = 0.197$  (0.247). Для данной пары изображений при анализе только границ для всех исследованных алгоритмов  $AEE$  наоборот снижается в 1.0–1.41 раза.

На паре изображений Venus количество объектов невелико; при этом они имеют сложную текстуру, а визуально количество границ объектов кажется большим (рис. 5, *б*). Однако согласно данным модельных перемещений (рис. 5, *в*) на изображении перемещаются лишь четыре крупных объекта. Результаты определения оптического потока в целом подобны предыдущей серии (Dimetraddon). Несмотря на это, судя по данным расчета параметра  $AEE$ , наименьшая ошибка характерна для CWFS —  $AEE = 0.363$  (0.339); чуть выше данный параметр для Simple Flow —  $0.385$  (0.418) (с учетом ошибочно выделенной белой области в верхнем правом углу), в то время как худший результат показал Classic NL, оптический поток опять-таки выглядит достаточно зашумленным:  $AEE = 0.450$  (0.410). Таким образом, для изображений стереопары использование всех трех алгоритмов дает близкие значения  $AEE$ . Для данной пары изображений ошибка определения оптического потока при анализе только границ по сравнению с рассмотрением всего изображения остается примерно постоянной.

Пара изображений Hydrangea содержит один объект со сложной текстурой на довольно однородном фоне. Из-за этого количество границ представляется достаточно большим (рис. 5, *б*). В результате на карте модельного оптического потока возникает много областей перекрытий (черные фрагменты на рис. 5, *в*). Сопоставление трех исследованных алгоритмов показало, что наименьшее значение  $AEE$  характерно для Classic NL —  $AEE = 0.410$  (0.207); при этом поле оптического потока внутри объекта снова выглядит достаточно “зашумленным” (рис. 5, *е*). Величина ошибки оказалась выше для CWFS —  $0.532$  (0.251); в этом случае внутри объекта смещения оказались весьма сглаженными (рис. 5, *д*). Подобная картина распределения оптического потока характерна и для Simple Flow, несмотря на то что величина  $AEE$  для него — максимальная —  $AEE = 0.608$  (0.309). В целом для данной пары изображений ошибка определения оптического потока при анализе только границ для всех исследованных алгоритмов возрастает в 1.97–2.12 раза.

Четыре последних анализируемых пары изображений являются искусственно синтезированными, что обуславливает ряд отличий в характере как построенных оптических потоков, так и в значениях  $AEE$ .

Обе серии Grove 2 и Grove 3 внешне выглядят подобно и содержат примерно идентичные объекты. Однако в серии Grove 2 мало объектов (листья — рис. 5, *б*, строка 5), в то время как на Grove 3 их значительно больше (линия 6). Это заметно и по картинкам модельных оптических потоков (рис. 5, *в*). Получено, что в такой ситуации более высокое визуальное качество построения оптического потока обеспечивает алгоритм Classic NL, что отражается и в минимальном значении  $AEE$ : для Grove 2 —  $0.308$  (0.270), для Grove 3 —  $0.982$  (0.808). Применение алгоритмов CWFS и Simple Flow обеспечивает получение чуть худших, но визуально подобных результатов (во втором случае смещения на краях объектов выглядят более четкими). Несмотря на это, ошибка построения оптического потока для развитого в работе алгоритма составила: для Grove 2 —  $0.364$  (0.296); для Grove 3 —  $0.985$  (0.761). Для алгоритма Simple Flow значение  $AEE$  выше:

Grove 2 — 0.484 (0.844), Grove 3 — 1.057 (0.844). На основании полученных данных можно констатировать, что при достаточно плотном расположении мелких объектов предложенный алгоритм CWFS не позволяет получить высококачественные результаты определения перемещений на их границах. Величина АЕЕ при анализе только границ для всех исследованных алгоритмов возрастает для серии Grove 2 в 1.14–1.35 раза, в то время как для Grove 3 она увеличивается примерно подобным образом — в 1.22–1.29 раза.

Еще две пары “синтетических” изображений также характеризуются значительным количеством объектов на них, которые имеют более правильную (прямоугольную) форму. Количество границ объектов также велико (рис. 5, б, строки 7, 8). С позиции же модельных оптических потоков в обоих случаях на изображениях наблюдается перемещение ограниченного количества объектов (рис. 5, в, строки 7, 8). Для серии Urban 2 визуальное лучшее качество построения оптического потока обеспечивает предложенный в работе алгоритм CWFS, в то время как для Urban 3 ближе к модельному оптическому потоку представляются результаты, полученные с помощью алгоритма Classic NL (несмотря на то, что опять-таки распределение достаточно “зашумленное”).

Сравнение количественных значений АЕЕ показало, что для серии Urban 2 результаты субъективного восприятия заметно отличаются от полученных численных оценок. Наименьшая величина АЕЕ при этом характерна для Simple Flow — 1.548 (0.928); несколько выше она для Classic NL — 1.583 (0.863); наибольшая величина ошибки получена как раз для CWFS — 2.139 (0.980). Для серии Urban 3 абсолютные значения ошибки получились значительно больше, а ранжирование алгоритмов поменялось: наименьшая величина АЕЕ теперь характерна для Classic NL — 1.191 (1.033); почти в три раза выше значение ошибки определения оптического потока для CWFS — 3.206 (1.713); еще выше значение АЕЕ для алгоритма Simple Flow — 4.295 (1.903).

Таким образом, для двух последних пар “синтетических” изображений при использовании алгоритмов Simple Flow и CWFS величина ошибки выросла в разы. Это объясняется тем, что описанные алгоритмические решения имеют подобную основу. По совокупности результатов обработки “синтетических” изображений алгоритм Classic NL следует считать более устойчивым и обеспечивающим меньшую ошибку определения оптического потока. В качестве наглядного проявления заметим, что для серии Urban 3 (строка 8) отчетливо видно фиолетовое пятно в центре изображения, отсутствующее на модельном потоке, а также при расчете с помощью Classic NL.

В целом ошибка определения оптического потока при анализе только границ по сравнению с рассмотрением полного изображения для всех исследованных алгоритмов для серии Urban 2 возрастает в 1.67–2.18 раза, в то время как для Urban 3 — в несколько меньшей степени — в 1.34–1.69 раза.

В рамках третьего эксперимента для двух серий изображений RubberWhale и Dimetrodon для предложенного в работе алгоритма CWFS оценивали ошибку вычисления оптического потока только вблизи границ при варьировании параметров его функционирования. В табл. 2 приведены результаты расчета АЕЕ при использовании подобранных в работе “рациональных” параметров, а также при отклонении от них. При подборе “рациональных” значений использовали также данные, приведенные в литературе.

Видно, что отклонение от “рациональных” значений параметров алгоритма CWFS (значения, приведенные до знака “/”), как правило, приводит к возрастанию АЕЕ. С другой стороны, в ряде случаев (для изображений RubberWhale) отклонение от “ра-

Т а б л и ц а 2. Средняя ошибка определения конечной точки АЕЕ при анализе смещений только на границах объектов при двух значениях параметров алгоритма CWFS

Величины параметров, используемых в алгоритме CWFS	Изображение типа Dimetrodon	Изображение типа RubberWhale
$\bar{m} = 7.7 / 3.3$	0.247 / 0.258	0.197 / 0.201
$\sigma_d = 22.5 / 45.5$	0.247 / 0.247	0.194 / 0.194
$\sigma_{fd} = 22.5 / 11$	0.247 / 0.248	0.194 / 0.192
$\sigma_c = 128 / 64$	0.247 / 0.247	0.194 / 0.194
$\sigma_{fc} = 36 / 18$	0.247 / 0.247	0.197 / 0.191
С учетом подобранных “рациональных” значений параметров алгоритма	0.247	0.194

ационального” значения параметров может сопровождаться получением меньшей интегральной ошибки определения оптического потока. Это означает, что их значения не являются универсальными для всех типов изображений, а зависят от величин перемещений, уровня зашумленности, количества объектов и плотности расположения их границ, контрастности изображений и т. п. По этой причине при их определении следует исходить из результатов анализа, полученных при обработке изображений различного типа (например, оценивая усредненную ошибку по всем типам изображений).

## Заключение

- В работе предложен и протестирован алгоритм вычисления оптического потока для пар изображений, ориентированный на повышение точности определения перемещений на границах объектов. В его основе лежит комбинация таких методов, как: а) блочное сопоставление; б) расчет устойчивой меры подобия блоков; в) анализ нескольких каналов информации; г) иерархический подход (пирамиды изображений); д) корректировка функции подобия блоков изображения на основе весов билатерального фильтра; е) субпиксельное уточнение смещений путем аппроксимации функции подобия блоков изображения вблизи максимума соответствующего распределения; ж) сглаживание вычисленного поля оптического потока.
- Результаты тестирования предложенного алгоритма в сопоставлении с описанными Classic NL и Simple Flow показали, что “в среднем” алгоритм CWFS обеспечивает меньшую величину определения оптического потока по сравнению с Simple Flow, однако уступает Classic NL: значение АЕЕ составляет 0.598 / 0.681 / 0.501 для всего изображения и 1.013 / 1.142 / 0.673 для учета только границ объектов соответственно.
- Предложенный алгоритм показал лучшие качественные и количественные результаты (минимальное значение АЕЕ) определения оптического потока при малых и средних величинах перемещений. Таким образом, он может быть рекомендован для применения в алгоритмах стереозрения (где взаимные перемещения объектов на изображении, как правило, невелики).
- При обработке “синтетических” изображений Grove применение алгоритмов Simple Flow позволяет получить ошибку АЕЕ построения оптического потока, близкую к таковой для Classic NL, в то время как для серий Urban ошибка выше в 1.5 раза. Таким образом, по совокупности результатов обработки алгоритм Classic NL сле-

дует считать более устойчивым и обеспечивающим меньшую ошибку определения оптического потока на “синтетических” изображениях.

- Последовательное изменение параметров алгоритма CWFS при обработке двух пар фотографий позволило определить их рациональные значения, обеспечивающие снижение АЕЕ. Однако достигаемое снижение ошибки определения оптического потока весьма невелико. Кроме того, “рациональные” значения, обеспечивающие снижение ошибки для одной пары изображений, могут определять ее повышение для другой пары. Поэтому их выбор должен проводиться индивидуальным образом.
- В качестве практической рекомендации отметим, что особенностью предложенного алгоритма является выполнение большей части вычислений в пределах локальных областей, что обуславливает возможность их распараллеливания. Это открывает значительную перспективу повышения быстродействия данного алгоритма с применением аппарата параллельных вычислений (например, с помощью графических процессоров), что будет являться предметом наших дальнейших исследований.

**Благодарности.** Работа выполнена в рамках программы фундаментальных исследований государственных академий наук (ПФИ ГАН) на 2013–2020 гг. по теме 23.1.3 “Научные основы диагностики предразрушения, оценки механического состояния и мониторинга многоуровневых структурно-неоднородных сред”.

## Список литературы / References

- [1] **Schreier, H., Orteu, J., Sutton, M.A.** Image correlation for shape, motion and deformation measurements. Basic concepts, theory and applications. N.Y.: Springer Sci., 2009. 321 p.
- [2] LaVision Techniques of digital image correlation. Available at: <http://www.lavision.de/en/techniques/dic-dvc/index.php> (accessed 25.06.2016).
- [3] **Fortun, D., Bouthemy, P., Kervrann, C.** Optical flow modeling and computation: A survey // Computer Vision and Image Understanding. 2015. Vol. 135. P. 1–21.
- [4] **Horn, B., Schunck, B.** Determining optical flow // Artificial Intelligence. 1981. Vol. 16. P. 185–203.
- [5] **Sundaram, N., Brox, T., Keutzer, K.** Dense point trajectories by GPU-accelerated large displacement optical flow // Lecture Notes in Comput. Sci. (Proc. ECCV, 2010.) 2010. Vol. 6311. P. 438–451.
- [6] **Baker, S., Black, M.J., Lewis, J., Roth, S., Scharstein, D., Szeliski, R.** A database and evaluation methodology for optical flow // Proc. of the Intern. Conf. on Comput. Vision (ICCV), 2007, Rio de Janeiro, Brazil. DC, USA: IEEE Comput. Soc., 2007. P. 1–8.
- [7] **Baker, S., Scharstein, D., Lewis, J., Roth, S., Black, M., Szeliski, R.** A database and evaluation methodology for optical flow // Computer Vision. 2011. Vol. 92, No. 1. P. 1–31.
- [8] **Barnes, C., Shechtman, E., Finkelstein, A., Goldman, D.** PatchMatch: a randomized correspondence algorithm for structural image editing // ACM Trans. on Graphics. (Proc. SIGGRAPH, Aug. 2009.) 2009. Vol. 28, No. 3. Article No. 24.
- [9] **Chen, Z., Jin, H., Lin, Z., Cohen, S., Wu, Y.** Large displacement optical flow with nearest neighbor field // Proc. of the Conf. on Comput. Vision and Pattern Recognition Workshops (CVPRW '13), June 23–28, 2013. Washington, DC, USA: IEEE Comput. Soc., 2013. P. 2443–2450.

- [10] **Kim, T., Lee, H., Lee, K.** Optical flow via locally adaptive fusion of complementary data costs // IEEE Intern. Conf. on Computer Vision. Dec. 2013. P. 3344–3351.
- [11] **Zimmer, H., Bruhn, A., Weickert, J., Valgaerts, L., Salgado, A., Rosenhahn, B., Seidel, H.-P.** Complementary optic flow // Lecture Notes in Comput. Sci. (Proc. EMMCVPR. Aug. 2009.) 2009. Vol. 5681. P. 207–220.
- [12] **Tao, M., Bai, J., Kohli, P., Paris, S.** SimpleFlow: A non-iterative, sublinear optical flow algorithm // Comput. Graph. Forum (Eurographics 2012). 2012. Vol. 31, No. 2. P. 345–353.
- [13] Itseez OpenCV. Available at: <http://www.opencv.org>. (accessed 25.06.2016).
- [14] **Werlberger, M., Pock, T., Bischof, H.** Motion estimation with non-local total variation regularization // Proc. of the Conf. on Comput. Vision and Pattern Recognition (CVPR), June 2010, San Francisco, CA, USA: IEEE, 2010. P. 2464–2471.
- [15] **Rosenberg, Y., Werman, M.** Representing local motion as a probability distribution matrix applied to object tracking // Proc. of the Conf. on Computer Vision and Pattern Recognition (CVPR '97), 17–19 June 1997, San Juan, PR, USA: IEEE Comput. Soc., 1997. P. 654–659.
- [16] **Гонсалес, Р., Вудс, Р.** Цифровая обработка изображений. М.: Техносфера, 2005. 1072 с.  
**Gonzalez, R.C., Woods, R.E.** Digital image processing. Upper Saddle River: Prentice Hall, 2002. 793 p.
- [17] **Anandan, A.** Computational framework and an algorithm for the measurement of visual motion // Intern. J. of Comput. Vision. 1989. Vol. 2, No. 3. P. 283–310.
- [18] **Kweon, I.S., Yoon, K.-J.** Adaptive support-weight approach for correspondence search // IEEE Trans. on Pattern Anal. and Machine Intelligence. 2006. Vol. 28, No. 4. P. 650–656.
- [19] **Debella-Gilo, M., Kaab, A.** Sub-pixel precision image matching for measuring surface displacements on mass movements using normalized cross-correlation // Remote Sensing of Environment. 2011. Vol. 115. P. 130–142.
- [20] **Xu, L., Jia, J., Matsushita, Y.** Motion detail preserving optical flow estimation // IEEE Trans. on Pattern Anal. and Machine Intelligence. 2012. Vol. 16, No. 9. P. 1744–1757.
- [21] **Sun, D., Roth, S., Black, M.J.** Secrets of optical flow estimation and their principles // The Twenty-Third IEEE Conf. on Comput. Vision and Pattern Recognition, CVPR 2010, 13–18 June, San Francisco, CA, USA: IEEE, 2010. P. 2432–2439. DOI:10.1109/CVPR.2010.5539939.

*Поступила в редакцию 24 июня 2016 г.,  
с доработки — 10 мая 2017 г.*

### **Determination of optical flow by complex algorithm on weighted full search**

PANIN, SERGEY V.<sup>1,2,\*</sup>, CHEMEZOV, VITALIY O.<sup>1</sup>, LYUBUTIN, PAVEL S.<sup>1</sup>

<sup>1</sup>Institute of Strength Physics and Materials Science SB RAS, Tomsk, 634055, Russia

<sup>2</sup>National Research Tomsk Polytechnic University, Tomsk, 634050, Russia

\*Corresponding author: Panin, Sergey V., e-mail: [svp@ispms.tsc.ru](mailto:svp@ispms.tsc.ru)

One of the important applied directions of machine vision is related to solving problems of experimental mechanics and particularly estimating three-dimensional deformation through analysis of image sequences or videostreams. The operating principle

for the systems of this kind is underlined by the technique of digital image correlation (DIC). However, the algorithms employed in the commercially available systems are not sufficiently accurate for determination of displacements at the edges of the objects. This results in difficulties in construction for a three-dimensional model of the working scene in 3D vision. In doing so, the problem related to the development of optical flow search algorithms in the three-dimensional computer vision systems is actual both in scientific and applied areas.

To improve the accuracy of determination for an optical flow, both at the edges of objects, as well as over the entire image it was suggested to combine the best approaches based on local analysis of images as well as “patch-match” technique. Local methods were chosen since they ensure the opportunity to perform parallel calculations.

In doing so, the motion estimation algorithm consists of a hierarchical search of weighted “patch-match” function with attracting more information channels obtained by preprocessing of input images followed by sub-pixel precision approximation at each hierarchical level and the final filtration of the obtained optical flow.

Test results of the proposed algorithm have demonstrated its competitiveness. The latter possessed the accuracy of the optical flow determination being comparative to the best competitors with a clear advantage to determine the optical flow at the edges of objects.

*Keywords:* image processing, optical flow, motion estimation, patch match, edges of objects, end point error, weighted block difference function, complex algorithm, subpixel accuracy.

**Acknowledgements.** The work was performed as part of the program of fundamental researches of the state academies of sciences (PFI SAS) for 2013–2020 23.1.3 on “Scientific bases of prefracture diagnostics, evaluation and monitoring of the mechanical state of heterogeneous media”.

*Received 24 June 2016*

*Received in revised form 10 May 2017*